

بسم الله الرحمن الرحيم

کلاس نکته و تست
نظریه زبان ها و ماشین ها

ابوالفضل گیلک

جلسه شانزدهم – فصل هشتم
خصوصیات بستاری زبان های مستقل از متن
تست 1 تا 16

حل تست های 15 سال اخیر

ارشد – دکتری

مهندسی کامپیوتر – علوم کامپیوتر

فصل 8

خصوصیات بستاری زبان‌های مستقل از متن

تست 1 تا 16

1- فرض کنیم L_1, L_2 و L_3 سه زبان مستقل از متن دلخواه باشند. آنگاه کدام مورد زیر درست است؟ (منظور از L' متمم مجموعه‌ای L است.) (مهندسی کامپیوتر - 1404)

- (1) زبان $L_1 \cdot (L_2 \cup L_3)$ حتماً مستقل از متن است.
 (2) زبان $L_1 \cdot (L_2 \cup L_3)$ حتماً مستقل از متن نیست.
 (3) زبان $L'_1 \cdot (L_2 \cup L_3)$ حتماً مستقل از متن است.
 (4) زبان $L'_1 \cdot (L_2 \cup L_3)$ حتماً مستقل از متن نیست.

2- کدام یک از مسائل زیر، در مورد یک اتوماتون پشته‌ای غیرقطعی (NPDA) دلخواه به نام M ، تصمیم‌پذیر (decidable) است؟ (مهندسی کامپیوتر - 1404)

- (1) زبان اتوماتون M ، توسط یک گرامر مستقل از متن نامبهم تشخیص داده می‌شود.
 (2) زبان اتوماتون M ، برابر زبان یک گرامر مستقل از متن دلخواه است.
 (3) زبان اتوماتون M ، دارای بی‌نهایت رشته است.
 (4) زبان اتوماتون M ، برابر Σ^* است.

3- کدام یک از زبان‌های زیر، منظم است؟ (مهندسی کامپیوتر - 1404)

- (1) $L = \{a^n b^m c^n d^{2n} \mid 3n + m > 5, n, m, k, t > 0\}$
 (2) $L = \{a^n b^t c^k d^t \mid n + 2t + k > 5, n, m, k, t > 0\}$
 (3) $L = \{a^n b^m c^k d^t \mid n + m + k + t > 5, n, m, k, t > 0\}$
 (4) $L = \{a^n b^m c^k d^m \mid n + 2m + k > 5, n, m, k, t > 0\}$

4- کدام مورد، درست است؟ (مهندسی کامپیوتر - 1403)

- (1) الگوریتم پویش (parsing) برای زبان‌های مستقل از متن، همیشه از مرتبه‌نمایی است و به فرم گرامر وابسته نیست.
 (2) اگر گرامر یک زبان مستقل از متن، به فرم نرمال چامسکی باشد، آنگاه می‌توان الگوریتم پویش (Parsing) با مرتبه خطی از طول رشته‌های آن زبان داشت.
 (3) اگر گرامر یک زبان مستقل از متن، گرامری ساده (s-grammar) باشد، آنگاه می‌توان الگوریتم پویش (parsing) با مرتبه خطی از طول رشته‌های آن زبان داشت.
 (4) اگر قوانین لامبدا (اپسیلون) (λ -Productions) را از یک گرامر یک زبان مستقل از متن حذف کنیم، آنگاه می‌توان الگوریتم پویش (parsing) با مرتبه خطی از طول رشته‌های آن زبان داشت.

5- کدام مورد، زبان تولید شده توسط گرامر زیر را توصیف می‌کند؟ (مهندسی کامپیوتر - 1403)

$$S \rightarrow aSc \mid aS \mid A$$

$$A \rightarrow bAcc \mid bA \mid bAc \mid \varepsilon$$

- (1) $\{a^m b^n c^p \mid 2m + 2n \geq p\}$
 (2) $\{a^m b^n c^p \mid m + 2n \geq p\}$
 (3) $\{a^m b^n c^p \mid m + n > p\}$
 (4) $\{a^m b^n c^p \mid 2m + 2n \geq p\}$

6- به ازای دو عدد صحیح و ثابت $q > 0$ و $r > 0$ کدام یک از زبان‌های زیر مستقل از متن (context-free) بوده ولی منظم (regular) نیست؟ (مهندسی کامپیوتر - 1402)

- (1) $L = \{1^{mq^2+rm} \mid m \geq 0\}$
 (2) $L = \{1^{mq^2+rm} \mid m \geq 0\}$
 (3) $L = \{1^{mq^2+rq^2} \mid m \geq 0\}$
 (4) $L = \{1^{mq+r} 0^{qr} \mid m \geq 0\}$

7- فرض کنید L_1 و L_2 زبان‌های مستقل از متن و R زبانی منظم باشند، کدام گزینه لزوماً زبانی مستقل از متن نخواهد بود؟ (مهندسی کامپیوتر - 1401)

- (1) $L_1 \cup L_2$ (2) $R - L_2$ (3) $L_1 - R$ (4) $L_1 L_2$

8- سه زبان L_1 ، L_2 و L_3 با تعاریف زیر مفروضند، کدام گزاره صحیح است؟ (مهندسی کامپیوتر - 1400)

$$L_1 = \{w0^n \mid w \in \{a,b\}^*, n_a(w) = n_b(w) = n, |w| = 2n\}$$

$$L_2 = \{w0^n \mid w \in \{a,b\}^*, |w| = n\}$$

$$L_3 = \{w0^n \mid w \in \{a,b\}^*, n_a(w) = n \text{ یا } |w| = n\}$$

(1) L_2 و L_3 هر دو نوع مستقل از متن قطعی هستند ولی L_1 از این نوع نیست.

(2) L_2 مستقل از متن قطعی است ولی L_1 مستقل از متن غیرقطعی است.

(3) L_2 مستقل از متن قطعی و L_3 مستقل از متن غیرقطعی است.

(4) هر سه زبان از نوع مستقل از متن هستند.

9- زبان‌های $L_1 = \{x\#y \mid x, y \in \{0,1\}^*, x \neq y\}$ و $L_2 = \{x\#y \mid x, y \in \{0,1\}^*, x = y\}$ را در نظر بگیرید. در این خصوص کدام مورد درست است؟

(مهندسی کامپیوتر - 99)

(1) L_1 و L_2 هر دو مستقل از متن‌اند.

(2) L_1 مستقل از متن است ولی L_2 مستقل از متن نیست.

(3) L_2 مستقل از متن است ولی L_1 مستقل از متن نیست.

(4) هیچ‌کدام از زبان‌های L_1 و L_2 مستقل از متن نیستند.

10- زبان‌های زیر را در نظر بگیرید:

(مهندسی کامپیوتر - 99)

$$L_1 = \{w_1cw_2 : w_1, w_2 \in \{a,b\}^*, w_1 \neq w_2\}$$

$$L_2 = \{w_1cw_2 : w_1, w_2 \in \{a,b\}^*, w_1 = w_2\}$$

کدام یک عبارت زیر صحیح است؟

(1) L_1 مستقل از متن نیست، L_2 مستقل از متن است، L_1 مکمل L_2 است.

(2) L_1 مستقل از متن نیست، L_2 مستقل از متن نیست، L_1 مکمل L_2 نیست.

(3) L_1 مستقل از متن است، L_2 مستقل از متن نیست، L_1 مکمل L_2 است.

(4) L_1 مستقل از متن است، L_2 مستقل از متن نیست، L_1 مکمل L_2 نیست.

11- زبان $L = \{yxzx^Ry^R : x, y, z \in \{0,1\}^+\}$ (علوم کامپیوتر - 94)

(2) حساس به متن (CS) نیست.

(1) یک ماشین پشته‌ای قطعی (DPDA) ندارد.

(4) منظم است.

(3) مستقل از متن است ولی منظم نیست.

12- زبان $L = \{a^n b^n c^n : n \geq 0\}$ و زبان‌های وابسته آن به شرح زیر مفروض است:

I. $h(L)$, $h(a) = a$, $h(b) = bb$, $h(c) = b$

II. L^R

III. L^*

IV. $h(L)$, $h(a) = a$, $h(b) = bb$, $h(c) = a$

(مهندسی کامپیوتر - 93)

کدام گزینه صحیح است؟

(1) I مستقل از متن است و بقیه مستقل از متن نیستند.

(2) I و III مستقل از متن و بقیه نیستند.

(3) I و IV مستقل از متن و بقیه نیستند.

(4) II مستقل از متن نیست و بقیه هستند.

13- زبان $L = \{xy : x, y \in \{a, b\}^*\}$ با اعمال چهار محدودیت به شرح زیر مفروض است:

I: $x = y$, II: $x \neq y$, III: $x = y^R$, IV: تعداد حروف x با تعداد حروف y برابر نیست:

(مهندسی کامپیوتر - 92)

کدام گزینه در مورد آن صحیح است؟

(1) L با محدودیت I و II مستقل از متن نیست و با III و IV هست.

(2) L با محدودیت I، II و IV مستقل از متن است و با III نیست.

(3) L با محدودیت I، III و IV مستقل از متن است و با II نیست.

(4) L با هر چهار محدودیت مستقل از متن است.

14- اگر $|w|_a$ تعداد حروف a در کلمه w باشد زبان $L \subseteq \{0,1\}^*$ را به صورت زیر در نظر بگیرید:

$$L = \{w : \exists n \geq 0, |w|_0 = 2^n, |w|_1 = 2^{n+1}\}$$

(علوم کامپیوتر - 89)

کدام گزینه صحیح نیست؟

(1) متمم زبان L یعنی L^c یک زبان مستقل از متن نیست.

(2) L مستقل از متن نیست ولی L^c مستقل از متن است.

(3) L یک زبان مستقل از متن نیست.

(4) هم L و هم L^c هر دو مستقل از متن نیستند.

15- اگر A و B دو زبان DCF (Deterministic Context Free) روی حروف الفبای Σ باشند و $\$ \notin \Sigma$ ، آن‌گاه کدام گزاره نادرست

(علوم کامپیوتر - 87)

است؟

(1) زبان $A \cup B \subseteq (\{\$ \} \cup \Sigma)^*$ یک زبان DCF است.

(2) $A \cup B$ لزوماً یک زبان DCF است.

(3) $\Sigma^* - A$ یک زبان DCF است.

(4) اگر B منظم باشد آن‌گاه $A \cap B$ یک زبان DCF است.

16- اگر $L = \{0^n 1^n 0^n : n \in \mathbb{N}\}$ آن‌گاه:

(علوم کامپیوتر - 86)

(1) L^c مستقل از متن است.

(2) L^c منظم است.

(3) L مستقل از متن است.

(4) L^c و L مستقل از متن هستند.

پاسخ‌ها

فصل 8: خصوصیات بستاری زبان‌های مستقل از متن

تست 1 تا 16

1- گزینه (1) صحیح است.

زبان‌های مستقل از متن تحت اجتماع و الحاق، بسته‌اند اما تحت مکمل بسته نیستند. بنابراین $L_1 \cdot (L_2 \cup L_3)$ مستقل از متن است. زبان $L'_1 \cdot (L_2 \cup L_3)$ ممکن است مستقل از متن باشد یا نباشد.

2- گزینه (3) صحیح است.

$L(M)$ مستقل از متن است. در زبان‌های مستقل از متن سه مساله معروف تصمیم‌پذیر عبارتند از: تهی یا ناتهی بودن؟ متناهی یا نامتناهی بودن، عضویت یا عدم عضویت w در $L(M)$ ؟ پس گزینه 3 تصمیم‌پذیر است.

3- گزینه (3) صحیح است.

در گزینه 1 باید تعداد d ها دو برابر تعداد a ها باشد. در گزینه 2 و گزینه 4 باید تعداد d ها برابر تعداد b ها باشد. پس این زبان‌ها منظم نیستند. در گزینه 3 بین توان‌ها چنین رابطه‌ای وجود ندارد، محدودیت $n+m+k+t > 5$ نیز به جز برای تعدادی متناهی از n, m, k, t برقرار است. این گزینه شامل تمام رشته‌های $a^+b^+c^+d^+$ به جز یک تعداد متناهی از آنهاست پس منظم است.

4- گزینه (3) صحیح است.

الگوریتم پویش (parsing) برای زبان‌های مستقل از متن کاملاً وابسته به فرم گرامر است. در حالت کلی الگوریتم پویش خسته کننده سطح به سطح برای $w \in L(G)$ متوقف می‌شود اما برای $w \notin L(G)$ ممکن است متوقف نشود و مرتبه زمانی‌اش نامتناهی شود. اگر با حذف قواعد λ و متغیرهای یکانی و غیرمفید گرامر را غیرانقباضی کنیم، جستجوی سطح به سطح، برای تمام رشته‌ها متوقف می‌شود و مرتبه زمانی نمایی $O(|P|^{2n+1})$ دارد که $(n=|w|)$. در صورتی که گرامر مورد نظر ساده باشد، مرتبه زمانی الگوریتم پویش، $O(n)$ (خطی) برحسب اندازه w است. اما همه زبان‌های مستقل از متن دارای گرامر ساده نیستند. برای هر زبان مستقل از متن L ، برای $L - \{\lambda\}$ گرامر نرمال چامسکی دارد و برای آن الگوریتم پویای (CYK) مرتبه $O(n^3)$ دارد.

5- گزینه (2) صحیح است.

از قاعده بازگشتی $S \rightarrow aSc$ ، شبه‌رشته a^nSc^n تولید می‌شود. با استفاده از قاعده $S \rightarrow aS$ شبه‌رشته $a^na^kSc^n$ تولید می‌شود. با $S \rightarrow A$ به شبه‌رشته $a^na^kAc^n$ می‌رسیم. قواعد $A \rightarrow bAcc \mid bA \mid bAc \mid \varepsilon$ نشان می‌دهند که متغیر A به رشته b^ic^j تبدیل می‌شود که $j \leq 2i$. با این توضیحات داریم:

$$L(G) = \{a^na^kb^ic^jc^n \mid j \leq 2i\} = \{a^{n+k}b^ic^{j+n} \mid j \leq 2i\}$$

$j \leq 2i$ و $n \leq n+k$ پس $j+n \leq 2i + (n+k)$. با نامگذاری توان‌ها داریم:

$$L(G) = \{a^mb^ic^p \mid p \leq 2i+m\}$$

6- گزینه (2) صحیح است.

برای نمونه $q=1$ و $r=1$ قرار دهیم:

گزینه 1 برابر است با $L = \{1^{2m}0 : m \geq 0\}$ که منظم است.

گزینه 2 برابر است با $L = \{1^{m+1}0^m : m \geq 0\}$ که مستقل از متن و نامنظم است.

گزینه 3 برابر است با $L = \{1^{m+1}0 : m \geq 0\}$ که منظم است.

گزینه 4 برابر است با $L = \{1^{m+1}0 : m \geq 0\}$ که منظم است.

7- گزینه (2) صحیح است.

زبان‌های مستقل از متن تحت اجتماع و الحاق بسته‌اند؛ همچنین تحت تفاضل منظم و اشتراک منظم بسته‌اند. اما گزینه ۲ لزوماً مستقل از متن نیست. در این گزینه با فرض $R = \Sigma^*$ داریم: $R - L_2 = \overline{L_2}$ و زبان‌های مستقل از متن تحت مکمل بسته نیستند.

8- گزینه (3) صحیح است.

در L_1 ، ماشین پشته‌ای توان محاسبه $|w| = 2n_a(w) \wedge n_a(w) = n_b(w)$ را ندارد این زبان مستقل از متن نیست اما حساس به متن است. در L_2 فقط باید طول w باید با تعداد صفرها برابر باشد. ماشین پشته‌ای توان پذیرش آن را دارد. در ضمن چون الفبای 0 در w ظاهر نمی‌شود، انتقال‌ها در این ماشین می‌توانند قطعی باشند. L_2 مستقل از متن قطعی است. L_3 اجتماع دو زبان مستقل از متن است پس مستقل از متن است اما با توجه به آنکه رشته‌ها دارای دو فرم مختلف هستند ماشین پشته‌ای قطعی ندارد.

9- گزینه (2) صحیح است.

(مشابه تست کنکور ارشد ۹۴). با توجه به شرط $x = y$ زبان L_2 مستقل از متن نیست. اما L_1 مستقل از متن است. این زبان را می‌توان به صورت اجتماع دو مجموعه نوشت که در یکی از آنها $|x| \neq |y|$ است و در دیگری $|x| = |y|$ است اما حداقل در یک حرف با هم تفاوت دارند. همچنین گرامر مستقل از متن زیر L_1 را تولید می‌کند. متغیرهای X, Y هرکدام، نوعی از اعضای L_1 را تولید می‌کنند.

$G : S \rightarrow X|Y$
 $X \rightarrow aXa|bXb|aXb|bXa|Z\#|Z$
 $Y \rightarrow AbZ|BaZ|Ab|Ba$
 $Z \rightarrow aZ|bZ|a|b$
 $A \rightarrow aAa|bAb|aAb|bAa|aZ\#|a\#$
 $B \rightarrow aBa|bBb|aBb|bBa|bZ\#|b\#$

10- گزینه (4) صحیح است.

دو زبان که مکمل هم باشند باید دو ویژگی داشته باشند: اشتراک آنها تهی و اجتماع آنها شامل تمام رشته‌ها باشد. اما در این مورد رشته‌هایی مانند a و aba به هیچکدام از زبان‌های L_1 و L_2 تعلق ندارند پس این دو زبان مکمل هم نیستند. گزینه‌های ۱ و ۳ رد می‌شوند.

L_2 با شرط عضویت $w_1 = w_2$ مستقل از متن نیست. ماشین پشته‌ای توان محاسبه شرط تساوی $w_1 = w_2$ را برای $w_1, w_2 \in \{a, b\}^*$ ندارد. اما L_1 با شرط $w_1 \neq w_2$ مستقل از متن است. L_1 را می‌توان به اجتماع دو زبان مستقل از متن تبدیل کرد:

$P = \{w_1cw_2 : w_1, w_2 \in \{a, b\}^*, |w_1| \neq |w_2|\}$
 $Q = \{w_1cw_2 : w_1, w_2 \in \{a, b\}^*, |w_1| = |w_2|, w_1 \neq w_2\}$

در گرامر زیر متغیرهای A, B هرکدام، نوعی از اعضای L_1 را تولید می‌کنند.

$G : S \rightarrow A|B$
 $A \rightarrow aAa|aAb|bAa|bAb|Zc|cZ$
 $Z \rightarrow aZ|bZ|a|b$
 $B \rightarrow XbE|YaE$
 $X \rightarrow aXa|aXb|bXa|bXb|aEc$
 $Y \rightarrow aYa|aYb|bYa|bYb|bEc$
 $E \rightarrow aE|bE|\lambda$

11- گزینه (4) صحیح است.

$y^R x^R = (xy)^R$ پس رشته‌های این زبان به فرم $xyz(xy)^R$ هستند. با فرض $w = xy$ داریم:

$$L = \{wzw^R : |w| \geq 2, |z| \geq 1\}$$

با کمی دقت، می‌توان عبارت منظم معادل با L را نوشت:

$$00(0+1)^+00+01(0+1)^+10+10(0+1)^+01+11(0+1)^+11$$

پس L منظم است در نتیجه مستقل از متن، قطعی و حساس به متن نیز هست.

12- گزینه (1) صحیح است.

زبان L نمونه معروف زبان‌های حساس به متن است که مستقل از متن نیست. وارون آن یعنی L^R و بستار ستاره آن یعنی L^* نیز حساس به متن هستند و مستقل از متن نیستند. در (I) تصویر همریختی L برابر $h(L) = \{a^n b^{3n} : n \geq 0\}$ است که مستقل از متن است. در (IV) تصویر همریختی L برابر $h(L) = \{a^n b^{2n} a^n : n \geq 0\}$ است که مستقل از متن نیست.

13- گزینه (3) صحیح است.

با اعمال محدودیت I داریم $L = \{ww : w \in \{a, b\}^*\}$ این زبان حساس به متن است ولی مستقل از متن نیست. با محدودیت II داریم $L = (a+b)^+$ که منظم است در نتیجه مستقل از متن نیز است. زبان L با محدودیت III مجموعه رشته‌های آینه‌ای به طول زوج است پس مستقل از متن غیرقطعی است. هر رشته از Σ^* به جز رشته تهی را می‌توان به شکل xy نوشت طوری که اندازه x مخالف اندازه y باشد؛ پس زبان L با محدودیت IV برابر Σ^+ است در نتیجه مستقل از متن است.

14- گزینه (2) صحیح است.

در واقع رشته‌های عضو L باید دو شرط داشته باشند: یکی آنکه $N_0(w) = 2^n$ و دیگر آنکه $N_1(w) = 2N_0(w)$. چون 2^n تابع غیرخطی است پس L و L^C مستقل از متن نیستند.

اثبات این مطلب با لم تزریق امکان پذیر است. برای استفاده از لم تزریق، برای هر عدد طبیعی n رشته به فرم $w = 0^N 1^{2N}$ را که $N = 2^n$ در نظر بگیریم. به وضوح $|w| = 3N \geq n$ این رشته را به صورت $w = xyvzu$ با لحاظ کردن ویژگیهای لم تزریق تجزیه کنیم:

$$x = 0^{m_1}, y = 0^{m_2}, v = 0^{m_3}, z = 0^{m_4}, u = 0^{N-m_1-m_2-m_3-m_4} 1^{2N}$$

برای هر $i \geq 0$ باید $w = xy^i v z^i u$ عضو L باشد اما چنین نیست. (مثلاً به $i = 2$ دقت کنید).

برای L^C نیز با لم تزریق می‌توان نشان داد مستقل از متن نیست.

15- گزینه (2) صحیح است.

زبان‌های مستقل از متن قطعی تحت اجتماع بسته نیستند. مثلاً $\{a^n b^n : n \geq 1\}, \{a^n b^{2n} : n \geq 1\}$ مستقل از متن قطعی هستند اما اجتماع آنها مستقل از متن غیرقطعی است. پس گزینه ۲ نادرست است. از طرفی اگر A و B دو زبان مستقل از متن قطعی باشند و یک وجه تمایز بین رشته‌های عضو A و B وجود داشته باشد که با مشاهده آن مشخص شود که این رشته از A آمده است یا از B آنگاه $A \cup B$ مستقل از متن قطعی است. برای مثال اگر تمام رشته‌های A را در یک نماد خاص مانند $\$$ الحاق کنیم آنگاه $A \cup B$ مستقل از متن قطعی است زیرا $\$A$ و B یک وجه تمایز دارند. با مشاهده اولین حرف از هر رشته معلوم است که آن رشته قرار است از کدام مجموعه بیاید. پس گزینه 1 درست است. گزینه 3 نیز درست است زیرا مستقل از متن‌های قطعی تحت مکمل بسته اند. همچنین تحت اشتراک منظم بسته اند پس گزینه 4 درست است.

16- گزینه (1) صحیح است.

به دلیل تساوی همزمان سه تا از توانها زبان L حساس به متن است اما مستقل از متن نیست. همچنین L منظم نیست. پس L^c نیز نمی‌تواند منظم باشد. پس گزینه 1 درست است.

راه دیگر: فرض کنیم منظور طراح از اعداد طبیعی، $N = \{0, 1, 2, \dots\}$ باشد. پس L^c برابر است با:

$$L^c = \{0^n 1^m 0^k : n, m, k \geq 0, n \neq m \text{ or } m \neq k\} \cup \overline{L(0^* 1^* 0^*)}$$

L^c مستقل از متن است ولی منظم نیست پس گزینه ۱ صحیح است.

پیروز و سربلند باشید.

پایان جلسه شانزدهم نکته و تست